DOCUMENT RESUME

ED 139 103 EA 009 518

AUTHOR Wilcox, Ray T.
TITLE A Comparison of Secondary School Teachers Judged Effective by Principals, Current Students, and Graduates.
PUB DATE Nov 76
NOTE 17p.; Paper presented at the Annual Meeting of the California Educational Research Association (55th, Burlingame, California, November 18-19, 1976)
AVAILABLE FROM Ray T. Wilcox, 148 McKay Building, Brigham Young University, Provo, Utah 84602

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Administrator Attitudes; Effective Teaching; High School Graduates; Principals; *Secondary Education; *Statistical Analysis; Student Attitudes; *Surveys; Teacher Attitudes; *Teacher Evaluation
IDENTIFIERS Utah

ABSTRACT

In five junior high schools and five senior high schools in Utah, the "voting method" was used to identify four groups of teachers, including (a) a group nominated by each principal as the most effective teachers in his school, (b) a group nominated by current students as their favorite teachers, (c) a group nominated by graduates as the teachers who taught them the most, and (d) a control group selected randomly from the remaining teachers. These different judges of effective teaching were found to nominate different teachers, with very low rank-order correlations between the top ten teachers in each group in each school. By using standard evaluative instruments, different judges were also found to favor a different group of teachers. Principals rated the teachers in group (a) the highest; the teachers in group (c) rated themselves the highest; and students rated the teachers in group (b) the highest. No judge rated the teachers in the control group the highest. This finding was seen as partially validating the results of the "voting method" as used in this study. (Author/JG)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED139103

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A COMPARISON OF SECONDARY SCHOOL TEACHERS JUDGED EFFECTIVE BY PRINCIPALS, CURRENT STUDENTS, AND GRADUATES

Ray T. Wilcox
Professor of Secondary Education
Brigham Young University

In five junior high schools and five senior high schools located in Utah, four groups of teachers were identified. (a) a group nominated by the principal as most effective; (b) a group nominated by the current students as their favorite teachers; (c) a group nominated by graduates as having taught them the most; and (d) a control group selected randomly from the remaining teachers.

The teachers in these four groups were evaluated by themselves, by their principals, and by their students using evaluation instruments adapted by the researcher. Mean scores for each group were compared to determine if the teachers of any particular group were rated as being most effective using standard evaluative procedures. The characteristics possessed by the teachers in each group were identified by use of a teacher self-description form created by the investigator.

This study was partly designed to investigate whether the "voting method" was a suitable technique for identifying the most effective teachers and to test the hypothesis that several "judges" of effective teaching would nominate different teachers as being the most effective. This study was also designed to provide a rationale which might explain why the teachers in the various groups appealed to different judges and to provide clues which might serve as a basis for further investigation of the question of who is the best judge of effective teaching.

Use of the "Voting Method" in Nominating the Most Effective Teachers.

Teachers were nominated by the students, graduates, and principals. The students were asked to complete a short questionnaire listing their three favorite

EA 009 518

teachers in order of how well they liked them. A first-ranked teacher was given three points, a second-ranked teacher two points, and a third-ranked teacher one point. A tabulation was made of the total number of points given each teacher in the school and the teachers placed in rank order. The graduates were asked to complete a short questionnaire which asked them to list the three teachers under whom they had learned the most in their junior or senior high school experience in order of how much they had been taught. The principals were asked to list their ten top teachers in order of their effectiveness.

The ten schools were of different sizes. An attempt was made to survey all of the students in attendance on a given day in the voting for favorite teacher. For the graduate sample, the same students in the senior high schools who were being asked about their favorite teachers were also polled about the teachers in junior high school under whom they had learned the most. If a junior high school fed a senior high school which was not participating in the study, then permission was requested to ask a large sample of high school students to complete just that part of the questionnaire which asked about their junior high school teachers. In the larger senior high schools participating in the study, an effort was made to poll by mail at least half of the graduates of the immediate past school year. In the smaller senior high schools, an effort was made to poll all of the graduates.

Rank-Order Correlation Coefficients Among the Three Groups.

Rank-order correlation coefficients between the principal's ranking and the students' ranking, between the principal's ranking and the graduates' ranking, and between the students' ranking and the graduates' ranking were computed for each of the five junior high schools and the five senior high schools.

It was found that, by using the "voting method," different judges were truly nominating different teachers. The rank-order correlation coefficients obtained were fairly low at the junior high school level (.24 for principal and students, .50 for principal and graduates, and -.02 for students and graduates) and very low at the senior high school level (-.09 for principal and students, .10 for principal and graduates, and .07 for students and graduates).

It seems a little unusual that the rank-order correlations were higher for the junior high schools than for the senior high schools. Perhaps this difference might be due to the fact that junior high schools are usually smaller than senior high schools and, consequently, the principals of the junior high schools had fewer teachers from whom to select the best teachers, thereby increasing the likelihood of also choosing those teachers whom the students and graduates favored. Another possible explanation might be because of having fewer teachers, the principals of the junior high schools get to know their teachers better than do the principals of the senior high schools.

At both the junior and senior school level, the rank-order correlations were highest for the rankings between the principals and graduates. This finding would seem to suggest that these judges were probably using the most similar criteria with which to rate the teachers. Since the correlations between the rankings of the principals and students were lower, it would seem that those qualities which help make a teacher a favorite with the students were not considered by principals as being qualities which make the teacher most effective in their eyes. Graduates and principals were seeing more nearly "eye-to-eye" than were current students and principals.

It is noticeable that graduates and current students were quite far apart in their judgements of teachers. Since they rated their teachers so differently, the question naturally arises as to whether students and graduates were using different evaluative criteria or whether the difference was due to how the voting was conducted. To help answer this question, in one junior high school and in one senior high school, both the students and graduates were asked, "Who are your favorite teachers?" and "Who taught you the most?"

It was found that the rank-order correlations improved dramatically when students and graduates were both responding to the same question about their teachers. Except for the comparisons with the 7th and 8th graders' "favorite" and "taught most" teachers at the junior high level, all of the correlation coefficients were reasonably high. The graduates agree most closely with the seniors at the senior high school level and with the 9th graders at the junior high school. The 12th graders, however, agree more closely with the 10th and 11th graders than do the 9th graders with the 7th and 8th graders. The discrepancy between "Favorite" teachers and "Taught Most" teachers appears to decrease as the students grow older. It seems probable that the perspective from which students judge their teachers changes considerably from the time they start in 7th and 8th grade and the time they graduate from high school.

Principals' Evaluations of Teachers in the Four Groups.

As an independent validity check on the "voting" procedures and to see which groups would rate well with principals, the teachers in all four groups were evaluated by the school principal. In completing his evaluation, the principal did not know the group identity of any teachers except, perhaps, those he himself had

nominated for group one. Naturally it would be expected that those teachers he had nominated himself would fare very well in his evaluations. It was found that the teachers in group one did fare the best as expected. The mean of 18.67 out of a total possible of 20.00 at the junior high school level and the mean of 18.95 at the senior high school level was the highest in both instances.

It might also be expected that the control group, since they were chosen from teachers not nominated by principals, students, or graduates, would be rated lowest and they were, for both the junior and senior high school levels. The mean score for the teachers in group four at the junior high school level was 14.87 and for the senior high school level was 14.23.

The combined means for the total sample were 18.84 for group one, 16.51 for group two, 17.27 for group three and 14.49 for group four. It is interesting to note that the principals rated the teachers in group three, those nominated by graduates, consistently higher than they did the teachers in group two, those nominated by current students. There were no exceptions at the junior high school level and only one exception at the senior high school level. This finding is in agreement with the rank-order correlations reported previously which found the principals agreeing more with graduates than with current students.

An analysis of variance using the MAD computer routine was completed to examine the following variables: L (level of school, junior or senior high school), S (school, with five different schools at each level), G (group, with four groups total), and R (rating, the first being the principals' score and the second the teachers' own self-rating score). This analysis found only two terms to be significantly different, G and GR. The F-ratios for G and GR were significant at the

1% level of confidence. Consequently, these terms were singled out for single degree-of-freedom contrasts. The first analysis combined the principals' ratings and the teachers' self-ratings and compared the adjusted mean (18.056) for group one with the adjusted mean (15.472) for group four (control group), the adjusted mean for group two (16.669) with the adjusted mean (15.472) for group four, and the adjusted mean for group three (17.046) with the adjusted mean (15.472) for group four. The F-ratios obtained were all significant at the 1% level of confidence. A further analysis indicated that much of the difference in the adjusted means was due to the principals' ratings even more than the teachers' self-ratings.

The principals' evaluations of the teachers in groups one, two, and three all differed significantly, at the 1% level of confidence, from the principals' evaluations of the teachers in group four, the control group. Furthermore, there were no siginificant differences in the evaluations given by junior and senior high school principals or by the principal of one school when compared with the principal of another.

In looking at the sub-scales, the teachers in group two were given higher ratings by the principals than the teachers in group three on only one scale, relations with students. On this one sub-scale, they were not only rated higher than the teachers in group three, they were even rated slightly higher than the teachers in group one. On all the other sub-scales, subject-matter competence, appropriateness of assignments and academic expectations, and overall classroom effectiveness, group-three teachers were rated higher than group-two teachers. Judging from the principals' perspective, it may be possible that the students are listing those teachers as their favorites who have developed good relationships with them regardless of other factors.

Teachers' Self-Evaluations in the Four Groups.

Each teacher was asked to evaluate himself using the same rating instrument given to the school principals. The teachers in group three rated themselves the highest with a mean score of 17.60; the teachers in group one the second highest with a mean score of 17.39; the teachers in group two the third highest with a mean score of 16.97; and the teachers in group four the lowest with a mean score of 16.42. Comparing these scores with the principals' evaluations, the teachers' evaluations were more homogenious and the rank ordering of the teachers was different, with groups one and three being interchanged.

As reported in the previous section, an analysis of variance was completed which analyzed both the principals' evaluations and the teachers' self-evaluations. Since G (group) and GR (group-rating interaction) were highly significant, a number of single degree-of-freedom contrasts were made to help interpret this finding. While the principals' evaluations tended to heighten the differences in scores among the four groups more than the teachers' self-evaluations, the latter were sufficiently varied to find two differences which were significant. The teachers in group three rated themselves significantly higher than the teachers in group four, the control group, at the 1% level of confidence. The teachers in group one rated themselves significantly higher than the teachers in group four at the 5% level of confidence. However, the teachers in group two did not rate themselves significantly higher than the teachers in group four. There were also no significiant differences in the self-evaluations of the teachers at the junior high school level when compared with those given by the teachers at the senior high school level and there were no significant differences in the self-evaluations of the teachers of one school when compared with those given by the teachers of another school.

It is noteworthy that the teachers in group one saw themselves as less competent than did the principals, and the teachers in group four, the control group, saw themselves as more competent than did the principals. The teachers in groups two and three also perceived themselves as slightly more competent than did the principals, although these differences were too small to be significant.

It also seems noteable that the teachers in group two rated themselves highest on the second subscale, relations with students, just as did the principals in their evaluations of these same teachers. The teachers' self-evaluation score of 4.73 on this subscale compares very closely with the principals' score of 4.62. Furthermore, group-two teachers rated themselves higher in this dimension of teaching competence than the teachers in the other three groups. This one area seems to be their strong point and it does not appear to have gone unrecognized.

Students' Evaluations of Teachers in the Four Groups.

All of the teachers in the sample were asked to allow the students in any two of their classes to evaluate them using a one-page, twenty-item evaluation form which the investigator adapted. Seventy-eight of the eighty-eight senior high school teachers, or 89%, agreed while only thirty-five of the sixty junior high school teachers, or 58%, agreed. It is clearly evident that the junior high school teachers were more hesitant to allow student evaluations of their teaching than were the senior high school teachers.

The mean scores obtained by the teachers at the junior high school level were as follows: Group 1, 3.85; Group 2, 4.01; Group 3, 3.91; and Group 4, 3.75. The mean scores obtained by the teachers at the senior high school level were as follows: Group 1, 4.24; Group 2, 4.42; Group 3, 4.29; and Group 4, 4.10.

It can be seen that the teachers in group two, those nominated as favorite teachers by the students, were in first place. Then came the teachers in group three, those nominated by the graduates to have taught them the most. Then came the teachers in group one, those nominated by the principal as the most effective teachers. And finally came the teachers in group four, those assigned to the control group. The same rank order was found at both the junior and senior high school levels, although the scores were noticeably higher for the senior high school teachers.

An analysis of variance using the MAD computer routine was completed to examine the following variables: L (level of school, junior or senior high school), S (school, with five different schools at each level), and G (group, with four groups total.) This analysis found all three variables, L, S, and G to be significant. L was already a single degree-of-freedom contrast and the investigator was not interested in learning which particular schools had students who rated their teachers higher than students in other schools. Consequently, only G was singled out for further single degree-of-freedom contrasts. These contrasts showed that the students rated the teachers in group two significantly higher than they rated the teachers in group four, the control group, at the 1% level of confidence. They also rated the teachers in group three significantly higher than they rated the teachers in group four, at the 5% level of confidence. But they failed to rate the teachers in group one significantly higher than they rated the teachers in group four.

Since L was significant at the 1% level of confidence, it is obvious that the senior high school teachers were evaluated more generously by their students than were the junior high school teachers. Judging from the reluctance shown by the junior high school teachers to participate in this phase of the research,

it would appear that the junior high school teachers are already aware of this tendency of their students to evaluate them more harshly and which may well have been a factor in the refusal of so many to participate in the students' evaluation.

Comparisons Among the Three Evaluations of Teachers.

It seems rather significant that each "judge" of effective teaching, the principals, the teachers themselves, and the students, gave preference to a different group of teachers when using objective evaluative criteria. This finding would appear to validate the previously reported low rank-order correlation coefficients found by using the "voting method" and the earlier conclusion that different judges of effective teaching truly do nominate different teachers.

Interestingly, the principals rated the teachers in group one the highest; the teachers in group three rated themselves the highest; and the students rated the teachers in group two the highest. The only clear agreement was that no judge rated the teachers in group four, the control group, the highest.

It is perhaps significant that, of these three "judges" of effective teaching, the principals, the teachers, and students, only one judge, the principals, tended to spread out the ratings very much from high to low.

Characteristics Possessed by Teachers in the Four Groups.

An attempt was made to identify some distinguishing characteristics possessed by teachers in the four groups. A five page, 160-item questionnaire (Teacher Self Description Form) was sent to each of the 148 teachers in the sample. Only two teachers, both at the senior high school level, failed to respond. Two teachers at the junior high school level also failed to respond but did so early in the study so there was time to select alternates who did respond. So, a total of 146 teachers, or 99% of the total sample, completed the questionnaire.

The results were summarized by group for each of the following areas: (1) Classroom Management and Discipline, (2) Motivational Techniques, (3) Teaching Methods, (4) Evaluation and Grading, (5) Value Education, (6) Personal Activities and Accomplishments, (7) Personal Beliefs, and (8) Personal Attributes. These summaries were used to provide a rationale which might explain why the teachers in the various groups appealed to the different judges of teaching effectiveness.

A Rationale for the Teachers in the Three Groups Which Explains Their Appealing to the Different Judges of Effective Teaching.

A clear rationale was found which explains why the teachers in group two would appeal to the students who listed them as their "favorite" teachers. When compared to the other teachers in the sample, these teachers reported themselves as more experimental, patient, gentle, pleasant, democratic, flexible, and informal than the others. When compared to what the other teachers did, they more strongly aimed to broaden student activities to include things other than reading, writing, and listening; tried to make their classes student-centered rather than teacher dominated; joked frequently with their students; knew the names of all of their students and greeted them outside of class; used a relaxed conversational style in their teaching; used a lot more student praise and encouragement than negative feedback and criticism; asked students for their opinions before making important class decisions; tried to make learning "fun" for their students; encouraged their students to help make the class rules; reinforced desirable student behavior with praise or other rewards but usually ignored undesirable behavior; used a few games and contests to motivate and interest their students; possessed a strong interest in sports; possessed awards and/or special

recognition in areas other than teaching; had a background of varied and extensive travel; tended to favor "open classrooms" over "traditional classrooms;" believed that teachers should try to bring about needed "social reforms" in society; believed that it will be possible someday to scientifically measure teaching effectiveness; believed that the teacher should be quite original in what is said and done in the classroom, and believed that educators should drastically change the traditional grading system. Most of these personal qualities and teaching behaviors would strongly appeal to the typical student. Add to this description the fact that principals evaluated these teachers so highly on sub-scale two, relations with students, and the fact that the teachers also rated themselves so highly in this same area. These evaluations and the above-mentioned comparative description of the teachers in group two provides an explanation as to why they were nominated as "favorite teachers" by the students.

Similarly, there also seems to be a clear rationale which explains why the teachers in group three would appeal to the graduates who listed them as the teachers who "taught them the most." When compared to the other teachers in the sample, these teachers reported themselves as more concise, fair, friendly, firm, confident, deliberate, systematic, outgoing, observant, interesting, loving, dependable, endowed with a high I.Q., endowed with a good sense of humor, a good public speaker, one who likes to teach, physically attractive, and endowed with a good personality than the others. When compared to what other teachers did, they more strongly restated student questions or comments to help clarify for the entire class; used the results of student tests to plan for future learning activities; fostered student inquiry by asking thought-provoking questions rather than recall questions; focused the attention of students on what they believed to be the critical

attributes of a topic or relationship; invited students to criticize the teachers' ideas and to express their own feelings and opinions; often allowed students a choice in selecting from a number of possible assignments; gave at least two student homework assignments per week; believed that they were generous in appraising the motives of their students; governed with as few rules as possible; insisted upon a work orientation in the classroom with a minimum of noise and confusion; commended students frequently and blamed them very seldom; gave descriptive, specific commendations rather than undescriptive, general commendations; called attention to the more interesting activities which students will be required to do during the weeks ahead; focused student attention on the lesson by using pictures, personal incidents, stories, news items, etc.; posed questions which are thought-provoking to stimulate student interest; gave mostly "objective-type" examinations; checked carefully to see that there is genuine consistency between their exams and the intended learning outcomes for their students; helped their students to build good self-images by setting realistic goals; tried to help their students behave in ways which are both moral and responsible; possessed a strong interest in reading; possessed a strong interest in music, painting, and the arts in general; possessed a strong interest in current events; exercised regularly to help maintain their physical fitness; had parents or other family members who are teachers; favored the use of "performance contracting" in grading students; and believed that teaching is more "science" than "art." Add to this description the fact that the teachers in group three evaluated themselves the highest of all the teachers on the teacher self-evaluation form. They have a good self-image and possess personal qualities and teaching behaviors which would help students to learn a good

deal and might serve to explain why the graduates nominated these teachers as those under whom they have "learned the most."

Since the teachers in group four were selected at random from the teachers not nominated by principals, current students, or graduates, it was not surprising to find them faring rather poorly in the comparison with the other teachers, but it was a real surprise to the investigator to find the teachers in group one also faring so poorly in the comparison. There doesn't seem to be any sort of clear rationale to explain why the teachers in this group appealed so strongly to the principals who nominated them as their best teachers. The most that can be said is that they were consistently "humble" in refusing to claim for themselves any of the descriptive "virtues" implied in the questionnaire. In comparing their responses to the responses of the other teachers, they did say they more strongly devoted themselves to tutoring students having difficulty; helped their students to develop talents and skills in their classes other than "academic;" had students coming to them for help and advice not directly related to the course; permitted students to sit where they choose; used their own silence to call attention to students who are talking when they should be quiet; permitted students to talk with each other for some of the time; gave A and B grades to more than 50% of their students; and believed that "teaching machines" should be available in the junior and senior high schools. This composite description from the questionnaire data does not seem to provide a clear rationale which might explain why the teachers in group one appeal so strongly to the principals. On the self-evaluation form, the teachers in group one rated themselves the second highest, just under the teachers in group three, so even the "extra humility" explanation is not consistent

with all of the facts. Further study is definitely needed to find a reasonable rationale for group-one teachers which might correspond with the rationales found for the teachers in groups two and three.

Using the Teacher Self-Description Form as a Second Self-Evaluation Form.

Since almost all of the questionnaire items were in the positive direction and reflected good teaching practices and desirable personal qualities, the researcher decided to use 156 of the 160 questionnaire items to serve as a second self-evaluation form. It was observed that group-three teachers rated themselves the highest on 82 items, group-two teachers rated themselves the highest on 37 items, group-one teachers rated themselves the highest on 18 items, and group-four teachers rated themselves the highest on 19 items. It was further observed that group-three teachers rated themselves the lowest on only 19 items, group-two teachers rated themselves the lowest on 44 items, group-one teachers rated themselves the lowest on 46 items, and group-four teachers rated themselves the lowest on 47 items.

Allowing the questionnaire items to serve in this manner as a second self-evaluation form, it is clear that the teachers in group three definitely viewed themselves as the most superior by a considerable margin, the teachers in group two as the next most superior, and the teachers in groups one and four as being tied approximately for the least superior. In fact, this analysis might well support a working hypothesis that, of the three judges of effective teaching, the principals, the students, and the graduates, the graduates are the best judge of effective teaching.

Summary of the Findings.

By using the "voting method," different "judges" of effective teaching were found to nominate different teachers with very low rank-order correlations being obtained among the top ten teachers in each of the three groups (excluding the control group) in each school.

By using standard evaluative instruments, different "judges" were also found to favor a different group of teachers. The principals rated the teachers in group one (those nominated by principals themselves) the highest; the teachers in group three (those nominated by the graduates) rated themselves the highest; and the students rated the teachers in group two (those nominated by the students) the highest. The only clear agreement was that no judge rated the teachers in group four (the control group) the highest.

These findings were seen as partially validating the results of the "voting method" as used in this study for nominating the most effective teachers.

By using the teacher self-description form, a fairly clear rationale was found to explain why the teachers in groups two and three appealed to the students and graduates. However, a similar rationale was not found to explain why group-one teachers appealed so strongly to the school principals.

Some evidence was also found to support a "tongue-in-cheek" hypothesis that, of the three "judges" in the study--principals, current students, and graduates, the graduates are the best judge of effective teaching. However, this evidence was based upon questionnaire data and not upon direct observation. Additional study, therefore, is needed to test such a hypothesis.